



UNITED STATES PATENT AND TRADEMARK OFFICE

fw
UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/653,039	08/29/2003	Ian M. Bennett	PHO 99004CIP	1543
23694	7590	03/02/2007	EXAMINER	
J. NICHOLAS GROSS, ATTORNEY 2030 ADDISON ST. SUITE 610 BERKELEY, CA 94704			LERNER, MARTIN	
			ART UNIT	PAPER NUMBER
			2626	

SHORTENED STATUTORY PERIOD OF RESPONSE	MAIL DATE	DELIVERY MODE
3 MONTHS	03/02/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

If NO period for reply is specified above, the maximum statutory period will apply and will expire 6 MONTHS from the mailing date of this communication.

Office Action Summary	Application No.	Applicant(s)
	10/653,039	BENNETT, IAN M.
	Examiner	Art Unit
	Martin Lerner	2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) Responsive to communication(s) filed on 02 February 2007.

2a) This action is FINAL. 2b) This action is non-final.

3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) Claim(s) 1 to 28 is/are pending in the application.

4a) Of the above claim(s) 16 to 21 is/are withdrawn from consideration.

5) Claim(s) _____ is/are allowed.

6) Claim(s) 1 to 15 and 22 to 28 is/are rejected.

7) Claim(s) _____ is/are objected to.

8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) The specification is objected to by the Examiner.

10) The drawing(s) filed on _____ is/are: a) accepted or b) objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a) All b) Some * c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. _____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) Notice of References Cited (PTO-892)
 2) Notice of Draftsperson's Patent Drawing Review (PTO-948)
 3) Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____

4) Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____

5) Notice of Informal Patent Application
 6) Other: _____

DETAILED ACTION

Election/Restrictions

Applicant's election without traverse of Group I, Claims 1 to 15 and 22 to 28, in the reply filed on 02 February 2007, is acknowledged.

Claims 16 to 21 are withdrawn from further consideration pursuant to 37 CFR 1.142(b) as being drawn to a nonelected invention, there being no allowable generic or linking claim. Election was made without traverse in the reply filed on 02 February 2007.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

Claims 1 to 3, 5, 9 to 10, 22, and 24 to 25 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Junqua et al.* in view of *Phillips et al.*

Concerning independent claims 1 and 22, *Junqua et al.* discloses a speech understanding system and method, comprising:

"a speech recognition engine for generating recognized words taken from an articulated speech utterance" – a spoken request and spoken information represented

as user speech at 30 is received by a speech recognizer 32; the spoken words are processed by the speech recognizer 32 and converted into text (column 3, lines 46 to 50: Figure 1);

“a natural language engine configured for linguistically processing said recognized words to generate search predicates for said articulated speech utterance” – natural language processor 34 includes local parser 36 and global parser 38 for further analyzing and understanding the semantic content of the digitized words provided by speech recognizer 32; local parser 36 examines the words using an LR grammar module 86 to determine if the word is recognized as a key word or a non-key word; when a word is recognized as a key word, the word is tagged with a data structure which represents the understood meaning of the word (column 4, lines 41 to 61: Figure 1); for a spoken request, “I would like to watch a movie tonight”, the key words are “watch”, “movie”, and “tonight” (column 5, lines 27 to 45); thus, the key words represent “search predicates”;

“a query formulation engine adapted to convert said recognized words and said search predicates into a structured query suitable for locating a set of one or more corresponding recognized matches for said articulated speech utterance” – dialogue manager 40 is assisted by a rule base 42 to perform a search within a program database; if the time key word slot 80 of a movie task frame 62 is filled, the dialogue manager 40 can search the program database 18 for all movies that begin at the requested time or during a time range; if the search produces more than a predetermined number of movies at the requested time, the dialogue manager 40 may

ask the user, "What type of movie would you like to watch?" (column 6, lines 25 to 43: Figure 1); thus, dialogue manager 40 is a "query formulation engine" for searching from "one or more corresponding recognized matches" in program database 18, given a "structured query" of filled key words slots, from "search predicates" as key words; "said natural language engine being further configured for linguistically processing said set of one or more corresponding recognized matches to determine a final match for said articulated speech utterance using both semantic decoding [and statistical based processing performed on said recognized words]" – natural language processor 34 includes a local parser 36 and a global parser 38; local parser 36 has the ability to analyze spoken grammatical expressions using an LR grammar module 86 to represent the understood meaning of the word; the examination is accomplished using a database of grammar structures (column 4, lines 41 to 65: Figure 1); natural language processor 34 is primarily responsible for analyzing the text stream and resolving the semantic content and meaning of the spoken request (column 3, lines 54 to 62: Figure 1).

Concerning independent claims 1 and 22, *Junqua et al.* discloses a natural language processor performing semantic decoding, but does not clearly say that the natural language processor performs statistical based processing. However, it is well known that speech recognition utilizes statistical processing to determine the probability that a word was spoken for a given speech utterance. *Phillips et al.* teaches dynamic semantic control of a speech recognition system that searches for keyword-value pairs with a dynamic semantic mechanism. Speech recognizer 102 outputs one or more

word strings that are the most probable words represented by the phonemes, ordered from the best to the worst, according to a probability that is created and stored in association with the word strings. Speech recognizer 102 uses dynamic semantic mechanism 112 to determine which words, from among the plurality of words, represent the semantics of the n-best words strings. (Column 5, Lines 43 to 61: Figure 1). Thus, speech recognizer 102 provides for both semantic decoding with dynamic semantic mechanism 112, and statistical based processing to produce n-best words strings with associated probability values. It is suggested that benefits may be realized from dynamic semantic models to provide greater accuracy compared to static word-based language models. (Column 4, Lines 60 to 65) It would have been obvious to one having ordinary skill in the art to perform statistical based processing as well as semantic decoding as taught by *Phillips et al.* in an apparatus and method for speech understanding of *Junqua et al.* for a purpose of providing greater recognition accuracy.

Concerning claims 2 and 3, *Junqua et al.* discloses local parser 36 tags key words for "a first level query", and global parser 38 places the key words into an appropriate slot 70 of an appropriate task frame 60 for "a second level query" (column 4, line 56 to column 5, line 26: Figure 1); implicitly, key words are being tagged by local parser 36 as they are being put into slots of a task frame by global parser 38 (Figure 2).

Concerning claim 5, *Junqua et al.* discloses that global parser 38 determines an appropriate task frame of what the user's desired action is, whether to watch a program, record a program, or inquire what programs are on (column 5, lines 17 to 45: Figures 1 and 2); a task frame defines an environment for a query, and is "a context parameter".

Concerning claims 9 and 10, *Junqua et al.* discloses a speech understanding system and method that operates in real time, implicitly; moreover, “notice” is taken that it is inherent that a speech recognition system produces results in less than 10 seconds for more than 100 potential matches, as a standard airline or train reservation system operating by speech recognition produces a result in less than 10 seconds for more than 100 potential cities.

Concerning claims 24 and 25, *Junqua et al.* discloses generating a preliminary query from key words, and a final query of tagged data structures that are placed into appropriate slots 70 of task frames 60 (column 5, lines 6 to 26: Figure 2); implicitly, linguistic processing by local parser 36 is occurring as global parser 38 puts each of the key words in slots.

Claim 4 is rejected under 35 U.S.C. 103(a) as being unpatentable over *Junqua et al.* in view of *Phillips et al.* as applied to claim 1 above, and further in view of *McDonough et al.*

Phillips et al. discloses semantic decoding based on a probability value for n-best strings, but omits calculating a term frequency based on a lexical distance between each word and one or more topic queries. However, *McDonough et al.* teaches topic discrimination for a speech recognition system, where one preferred method employs a Kullback-Liebler distance measure, providing a measure of dissimilarity of the occurrence patterns of an event for a given topic as opposed to all other topics. (Column 11, Lines 40 to 60) It is suggested that improved speech recognition can be

achieved if a potential topic can be detected for a set of potential speech events. (Column 3, Line 63 to Column 4, Line 24) It would have been obvious to one having ordinary skill in the art to calculate a term frequency based on a lexical distance between words and one or more topic queries as taught by *McDonough et al.* in a semantic speech recognition system of *Phillips et al.* for a purpose of improving speech recognition by topic discrimination.

Claims 6, 11 to 13, and 27 to 28 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Junqua et al.* in view of *Phillips et al.* as applied to claims 1 and 22 above, and further in view of *Barclay et al.*

Junqua et al. does not expressly disclose placing a speech query recognition system on a server computer, so that the speech recognition is distributed across a client-server architecture to reduce transmission latencies, multiple servers, and controlling a web page. However, distributed speech recognition in a client-server architecture is well known. Specifically, *Barclay et al.* teaches a client-server speech recognizer, where processing capabilities are distributed between the client and the server. (Abstract) A client digitizes speech, extracts features, and quantizes the features, and a server performs speech recognition and natural language understanding. Latency is reduced because lower bandwidths are required, as less data needs to be communicated between the client and the server. (Column 4, Lines 1 to 16) Speech recognition capabilities may be incorporated into a World-Wide-Web browser (column 8, lines 36 to 64: Figure 4), and a general architecture distributes

between one client and a plurality of servers (column 9, lines 31 to 42: Figure 6). An objective is to process speech with large vocabularies and grammars in real time with a client computer being a laptop. (Column 4, Lines 10 to 16) It would have been obvious to one having ordinary skill in the art to incorporate the client/server architecture for distributed speech recognition of *Barclay et al.* into a speech understanding system and method of *Junqua et al.* for a purpose of processing speech with large vocabularies and grammars in real time on a laptop.

Claims 7, 8, 14, 23, and 26 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Junqua et al.* in view of *Phillips et al.* as applied to claims 1 and 22 above, and further in view of *Appelt et al.* ('026).

Junqua et al. discloses natural language processing to determine semantic content, but omits determining noun-phrases to compare and to provide a final match; discloses searching with key words, but omits SQL search predicates; and discloses a speech synthesizer 44, but omits providing a final match in an audible form. However, *Appelt et al.* ('026) teaches information retrieval by natural language querying, where noun groups and noun phrases are utilized. (Column 7, Line 61 to Column 8, Line 29; Column 9, Lines 28 to 51) Responses to natural language queries are provided by converting a response to speech using a text-to-speech unit (column 2, lines 50 to 51), and a query is converted into an SQL query (column 6, lines 13 to 26). An objective is to provide search results to users in a timely fashion through natural language to support accurate and fast searches from multimedia sources of information. (Column 4,

Lines 22 to 39) It would have been obvious to one having ordinary skill in the art to provide the features of determining noun phrases, SQL search predicates, and audible answers as taught by *Appelt et al.* ('026) in a speech understanding system and method of *Junqua et al.* for a purpose of providing fast searches from multimedia sources of information.

Claim 15 is rejected under 35 U.S.C. 103(a) as being unpatentable over *Junqua et al.* in view of *Phillips et al.* as applied to claim 1 above, and further in view of *Agarwal et al.* ('196).

Junqua et al. omits a relational database that is updated asynchronously to reduce retrieval latency. However, *Agarwal et al.* ('196) teaches that it is common for relational databases to be updated in an asynchronous manner to avoid the inefficiencies of re-reading records. It would have been obvious to one having ordinary skill in the art to asynchronously update a relational database as taught by *Agarwal et al.* ('196) to search multimedia databases of *Junqua et al.* for a purpose of avoiding inefficiencies of re-reading records.

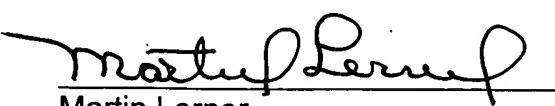
Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Martin Lerner whose telephone number is (571) 272-7608. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

ML
2/27/07



Martin Lerner
Examiner
Group Art Unit 2626